

**MODEL STATISTIK UNTUK DATA PENCEMARAN UDARA
OLEH OZON EKSTRIM DI PEKANBARU**

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Sains
pada Jurusan Matematika

Oleh :

AGUS DIANTORO
10854003918



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2013**

MODEL STATISTIK UNTUK DATA PENCEMARAN UDARA OLEH OZON EKSTRIM DI PEKANBARU

AGUS DIANTORO
10854003918

Tanggal Sidang : 28 Oktober 2013
Tanggal Wisuda : Maret 2014

Jurusan Matematika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau
Jl. HR. Soebrantas No.155 Pekanbaru

ABSTRAK

Tugas akhir ini untuk data ozon ekstrim, adapun distribusi yang digunakan yaitu distribusi *Generalized Extreme Value* dan distribusi Pareto. Kedua distribusi ini diterapkan pada data pencemaran udara oleh ozon ekstrim di Pekanbaru. Metode yang digunakan dalam mengestimasi parameter adalah *maksimum likelihood* untuk distribusi Pareto dan *probability weight moment* untuk distribusi *GEV*. Selanjutnya dilakukan uji kebaikan atau uji *goodness of fit* yaitu uji *Kolmogorov Smirnov* dan uji *Anderson Darling*. Dari hasil kedua uji yang digunakan diperoleh hasil bahwa distribusi *GEV* lebih baik dibandingkan dengan distribusi Pareto, karena menghasilkan nilai uji yang minimum dan bentuk kurva yang lebih mendekati kurva normal.

Katakunci : Data pencemaran udara oleh ozon ekstrim, distribusi *GEV*, distribusi Pareto, maksimum *likelihood*, *probability weight moment*.

***STATISTICAL MODEL FOR AIR POLLUTION DATA BY
OZONE EXTREME PEKANBARU***

***AGUS DIANTORO
10854003918***

***Date of Final Exam : 28 Oktober 2013
Date of Graduation Ceremony : March 2014***

*Departement of Mathematics
Faculty of Science and Technology
State Islamic University of Sultan Syarif Kasim Riau
HR. Soebrantas Street No.155 Pekanbaru*

ABSTRACT

This paper is for extreme ozone data, while the distribution used is Generalized Extreme Value distribution (GEV) and Pareto distribution. Both distributions are applied to the data of extreme air pollution by ozone in Pekanbaru. The method is used to estimate the parameters is maximum likelihood for the Pareto distribution and the probability weight moment for GEV distribution. Then, goodness of fit test is using the Kolmogorov Smirnov and Anderson Darling test. From the results of the two trials that used, the obtained results is that the GEV distribution is better than the Pareto distribution, since it produces the minimum value of the test and the shape of the curve is closer to the normal curve.

Keywords: *data of extreme air pollution by ozone, distribusi GEV, distribusi Pareto, maksimum likelihood, probability weight moment.*

KATA PENGANTAR

Alhamdulillah rabbil'alamin, segala puji bagi Allah SWT yang senantiasa melimpahkan rahmat dan taufik serta hidayah-Nya sehingga penulis dapat menyelesaikan tugas akhir ini seperti yang diharapkan. Adapun tugas akhir yang diselesaikan oleh penulis yaitu dengan judul "**Model Statistik untuk Data Pencemaran Udara oleh Ozon Ekstrim di Pekanbaru**". Tugas akhir ini merupakan salah satu syarat untuk memperoleh gelar sarjana sains (S1). Selanjutnya sebagai umat nabi Muhammad SAW, salawat serta salam harus selalu tercurahkan kepada Nabi Muhammad SAW karena atas jasa dan kepedulian yang diberikan kepada kita dengan membawa petunjuk bagi seluruh umat manusia.

Dalam penyusunan dan penyelesaian tugas akhir ini penulis tidak terlepas dari bantuan berbagai pihak, baik langsung maupun tidak langsung. Untuk itu sudah sepantasnya penulis mengucapkan terimakasih yang tak terhingga kepada kedua orang tua tercinta, ayah (Ronggo Warsito) dan ibu (Ponirah) yang tidak pernah lelah dan tiada henti melimpahkan kasih sayang, perhatian, motivasi yang membuat penulis mampu untuk terus dan terus melangkah, pelajaran hidup, juga materi yang tak mungkin bisa terbalas. Jasa-jasamu kan selalu kukenang hingga akhir hayatku dan semoga Allah menjadikan jasa-jasamu sebagai amalan soleh, Amin. Selanjutnya ucapan terimakasih kepada :

1. Bapak Prof. DR. H. M. Nazir, M.A selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Ibu Dra. Hj. Yenita Morena, M.Si selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Ibu Sri Basriati, M.Sc selaku Ketua Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Ibu Ari Pani Desvina, M.Sc selaku Pembimbing yang telah menjadi orang terpenting dalam penulisan tugas akhir ini dan banyak membantu, mendukung, mengarahkan dan membimbing penulis dalam penulisan tugas akhir ini.

5. Ibu Rahmadeni, M.Si selaku Penguji I dan yang telah memberikan kritikan dan saran yang membangun sehingga tugas akhir ini selesai.
6. Bapak Nilwan Andiraja, M.Sc selaku Penguji II yang telah memberikan kritikan dan saran sehingga tugas akhir dapat selesai.
7. Bapak Rado Yendra, M.Sc yang telah memberikan masukan dan memberikan jalan dalam memahami teori-teori dari tugas akhir yang akhirnya berhasil diselesaikan oleh penulis.
8. Bapak M. Marizal, M.Sc yang memberikan masukan berupa solusi yang bisa membantu penulis menyelesaikan skripsi ini.
9. Dosen jurusan matematika yang telah banyak membantu dalam penyelesaian Tugas akhir ini.
10. Kawan-kawan jurusan matematika angkatan 2008 yang semuanya ikut memberikan motivasi sehingga penulis berhasil menyelesaikan tugas akhir ini sesuai yang diharapkan.

Dalam penyusunan dan penulisan tugas akhir ini penulis telah berusaha semaksimal mungkin untuk menghindari kesalahan. Tapi tak ada manusia yang sempurna di dunia ini. Akhirnya penulis mengharapkan kepada pembaca tugas akhir ini agar memberikan saran dan kritik yang bersifat konstruktif. Semoga tugas akhir ini dapat memberikan konstribusi yang bermanfaat. Amin...

Pekanbaru, 22 Desember 2013

Penulis

DAFTAR ISI

JUDUL	Halaman
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iv
LEMBAR PERNYATAAN	v
LEMBAR PERSEMBAHAN	vi
ABSTRAK.....	viii
<i>ABSTRACT</i>	ix
KATA PENGANTAR	x
DAFTAR ISI.....	xii
DAFTAR SIMBOL	xiv
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xvi
DAFTAR SINGKATAN	xvii
DAFTAR LAMPIRAN.....	xviii
 BAB I PENDAHULUAN	
1.1 Latar Belakang Masalah.....	I-1
1.2 Rumusan Masalah	I-2
1.3 Batasan Masalah	I-2
1.4 Tujuan Penelitian	I-3
1.5 Manfaat Penulisan.....	I-3
1.6 Sistematika Penulisan	I-4
 BAB II LANDASAN TEORI	
2.1 Data Pencilan dan Data ekstrim	II-1
2.2 Ozon	II-1
2.3 Pencemaran Udara	II-2
2.4 Distribusi Peluang.....	II-2
2.5 Nilai Harapan Distribusi Peluang	II-3

2.6	Variansi Distribusi Peluang	II-4
2.7	Distribusi Pareto.....	II-5
2.8	Distribusi <i>Generalized Extreme Value (GEV)</i>	II-8
2.9	Estimasi Parameter.....	II-9
2.10	Fungsi <i>Likelihood</i>	II-9
2.10.1	Estimasi Maksimum <i>Likelihood</i>	II-10
2.10.2	Estimasi Parameter Menggunakan MLE.....	II-10
2.10.3	Estimasi <i>Probability Weight Moment</i>	II-12
2.10.4	Estimasi Parameter dengan PWM.....	II-13
2.11	Uji Kebaikan (<i>Goodness Of Fit</i>)	II-14
2.11.1	Uji Kolmogorov-Smirnov	II-15
2.11.2	Uji Anderson Darling	II-15
2.12	Penelitian-penelitian yang Terkait dengan Model	II-16
BAB III METODOLOGI PENELITIAN		
BAB IV ANALISA DAN PEMBAHASAN		
4.1	Deskriptif Statistik Data	IV-1
4.2	Menentukan Nilai Parameter Data Ozon Ekstrim.....	IV-2
4.3	Model Distribusi Untuk Data Pencemaran Dara Oleh Ozon Ekstrim di Pekanbaru	IV-4
4.4	Uji Kebaikan Model (<i>Goodness of Fit</i>).....	IV-6
4.4.1	Uji Kolmogorov Smirnov.....	IV-6
4.4.2	Uji Anderson Darling	IV-9
BAB V KESIMPULAN DAN SARAN		
5.1	Kesimpulan.....	V-1
5.2	Saran.....	V-1
DAFTAR PUSTAKA		
DAFTAR RIWAYAT HIDUP.....		

BAB I

PENDAHULUAN

1.1 Latar Belakang

Udara adalah salah satu sumber daya alam yang sangat penting bagi kebutuhan hidup manusia, hewan dan tumbuhan dalam mempertahankan hidupnya. Oleh karenanya, udara perlu dijaga kebersihannya, melalui pemantauan, pengaturan dan pengendalian polutannya sehingga tidak melampaui batas yang masih diperkenankan bagi kehidupan. Komposisi udara sendiri terdiri dari berbagai senyawa gas dan partikel pencemar yang tersebar bebas di udara. Senyawa gas yang tersebar diudara meliputi CO, NO_x, ozon (BMKG, 2012).

Pencemaran udara merupakan salah satu faktor penting yang perlu diperhatikan dan ditanggulangi dalam mencapai pembangunan berwawasan lingkungan. Sesuai dengan pembangunan nasional yaitu pembangunan manusia Indonesia seutuhnya dan pembangunan masyarakat Indonesia. Kajian tentang pencemaran udara yaitu gabungan antara udara atau atmosfir sebagai sumber daya alam dengan kepentingan manusia seperti kesehatan, keselamatan, kesejahteraan dan kenyamanan (K4). Supaya K4 tercapai maka perlu dijaga keselarasan, keserasian, kesetimbangan dan kesatuan tujuan yang utuh dalam setiap kegiatan pembangunan (Teknik Lingkungan ITB,2009).

Sumber pencemaran udara terdiri dari pembakaran, aktifitas kendaraan bermotor dan partikel yang tersebar di udara. Sumber pencemaran ini mengakibatkan aktivitas polutan ozon bisa berdampak buruk pada manusia, hewan dan tumbuhan seperti pemanasan global, hujan asam dan gangguan pernapasan. Sehubungan dengan itu perlu adanya analisis pada data ozon yang sesuai dan dapat digunakan untuk meramalkan kondisi ozon untuk masa yang akan datang. Memantau tingkat konsentrasi ozon adalah langkah yang baik untuk mengurangi tingginya pencemaran. Konsentrasi ozon yang tidak melebihi tingkat standar 0,10 ppm dianggap aman dan tidak membahayakan manusia. Oleh karena itu menganalisis konsentrasi yang melebihi tingkat aman tersebut adalah tujuan dari penulisan tugas akhir ini. Konsentrasi pencemar yang melebihi standar aman dan bernilai maksimum disebut sebagai nilai ekstrim (Bisono, 2011).

Secara alamiah ozon bisa terdapat di mana-mana dalam atmosfer. Ozon adalah senyawa yang terdiri dari 3 atom oksigen (O_3). Jika dibandingkan dengan gas oksigen yang hanya senyawa 2 atom oksigen (O_2) maka jelas ozon memiliki daya oksidasi yang luar biasa. Ozon berbentuk gas berwarna biru yang daya oksidasinya sangat tinggi. Ozon juga termasuk disinfektan yang lebih baik dari pada klor. Sifat interaksinya dengan sinar UV menjadi bagian terpenting dalam fungsinya sebagai perisai di planet bumi. Ozon sangat mudah menyerap sinar UV secara langsung yang berkisar diantara 240-280 nm, kemudian pecah menjadi 1 atom oksigen (O) dan 1 molekul gas oksigen (O_2) yang tersebar di udara (Sugata Pikatan, 1992).

Berdasarkan keterangan tersebut dan penelitian terkait seperti Anita Rahayu yaitu "*Estimasi Parameter Distribusi Generalized Extreme Value*" dan Wahyudi "*Identifikasi Curah Hujan Ekstrem di Kabupaten Ngawi Menggunakan Generalized Extreme Value dan Generalized Pareto Distribution*" yang meneliti pada data ekstrim, maka penulis tertarik untuk mencari distribusi nilai ekstrim yang terbaik dengan judul tugas akhir yaitu **"Model Statistik untuk Data Pencemaran Udara oleh Ozon Ekstrem di Pekanbaru"**.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang tersebut, maka rumusan masalah pada penelitian ini adalah :

1. Bagaimana menerapkan distribusi Pareto dan distribusi GEV, untuk memodelkan data pencemaran udara oleh ozon ekstrim di Pekanbaru Tahun 2009-2012
2. Bagaimana menentukan distribusi yang terbaik antara distribusi Pareto dan distribusi GEV untuk memodelkan data pencemaran udara oleh ozon ekstrim di Pekanbaru Tahun 2009-2012

1.3 BATASAN MASALAH

Untuk mencegah meluasnya permasalahan yang ada dan agar lebih terarah, maka dilakukan pembatasan masalah yaitu :

1. Data yang digunakan yaitu data pencemaran udara oleh ozon ekstrim harian di Pekanbaru dari Tahun 2009-2012
2. Distribusi yang digunakan adalah distribusi Pareto dan *Generalized Ekstrim Value (GEV)*.
3. Menentukan distribusi terbaik dari dua distribusi yang digunakan dilakukan dengan uji *goodness of fit* menggunakan uji statistik Anderson Darling dan Kolmogorov Smirnov.

1.4 TUJUAN PENELITIAN

Tujuan penelitian ini adalah:

1. Menerapkan distribusi Pareto dan distribusi *Generalized Ekstrim Value (GEV)*, untuk memodelkan data pencemaran udara oleh ozon ekstrim di Pekanbaru Tahun 2009-2012
2. Menentukan distribusi yang terbaik antara distribusi Pareto dan distribusi GEV, untuk memodelkan data pencemaran udara oleh ozon ekstrim di Pekanbaru Tahun 2009-2012

1.5 MANFAAT PENELITIAN

Adapun manfaat dalam penelitian ini adalah :

1. Bagi Penulis
Mendapatkan model distribusi yang terbaik untuk data pencemaran udara ozon ekstrim di Pekanbaru Tahun 2009-2012
2. Bagi Lembaga Pendidikan
Sebagai sarana informasi bagi pembaca dan sebagai bahan referensi bagi pihak yang membutuhkan.
3. Bagi Badan Lingkungan Hidup Kota Pekanbaru.
Memberikan gambaran tentang model distribusi yang terbaik untuk data pencemaran udara oleh ozon ekstrim di Pekanbaru.

1.6 SISTEMATIKA PENULISAN

Sistematika dalam pembuatan tulisan ini mencakup lima bab yaitu :

BAB I Pendahuluan

Bab ini berisi latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II Landasan Teori

Bab ini menjelaskan teori-teori nilai ekstrim, distribusi Pareto dan distribusi GEV.

BAB III Metodologi Penelitian

Bab ini berisikan langkah-langkah atau prosedur untuk menganalisis dan memodelkan nilai ekstrim dari data pencemaran udara oleh ozon ekstrim di Pekanbaru sepanjang Tahun 2009-2012.

BAB IV Pembahasan dan Analisa

Bab ini membahas tentang hasil yang diperoleh pada analisis data pencemaran udara oleh ozon ekstrim dan model terbaik yang diperoleh sehingga dapat digunakan sebagai gambaran kondisi untuk ozon ekstrim.

BAB V Penutup

Bab ini berisikan kesimpulan dan saran.

BAB II

LANDASAN TEORI

2.1 Data Ekstrim dan Data Pencilan

Identifikasi data merupakan bagian yang sangat penting dalam penelitian statistik. Karakteristik data yang sesuai akan membentuk model yang baik dan sesuai. Data yang sesuai dengan penelitian akan memberikan hasil yang lebih akurat pada penelitian yang dilakukan.

Data ekstrim adalah data yang bernilai maksimum atau minimum dari dalam selang waktu tertentu seperti dalam skala harian, bulanan, dan tahunan. Data yang bernilai maksimum disebut data ekstrim maksimum, dan data yang bernilai minimum disebut sebagai data ekstrim minimum.

Data pencilan adalah pengamatan ekstrim dan merupakan titik data yang mengalami penyimpangan cukup jauh dari seluruh pengamatan data (Anna Fauziyah, 2012).

2.2 Ozon

Ozon adalah komponen atmosfer yang diproduksi oleh proses fotokimia, yaitu suatu proses kimia yang membutuhkan sinar matahari untuk mengoksidasi komponen-komponen yang tidak langsung dioksidasi oleh oksigen. Senyawa yang terbentuk merupakan bahan pencemar sekunder yang diproduksi dari interaksi antara bahan pencemar primer dengan sinar matahari. Hidrokarbon merupakan komponen yang berperan dalam produksi oksidan fotokimia. Reaksi ini juga melibatkan siklus fotolitik NO_2 . Polutan sekunder yang dihasilkan dari reaksi hidrokarbon dalam siklus ini adalah Ozon dan Peroksiasetilnitrat (PAN). Karena Ozon merupakan senyawa yang dominan dari oksidan fotokimia ini, yaitu mencakup kira-kira 98% volume, maka hasil pemantauan udara ambien dinyatakan sebagai kadar Ozon (Soedomo, 2001).

Ozon dapat ditemukan di setiap tempat dimana terdapat oksida nitrogen dan hidrokarbon yang berinteraksi di bawah radiasi sinar matahari. Ozon berbahaya bagi tumbuh-tumbuhan, karena dapat mengganggu proses fotosintesis. Sedangkan

dampak terhadap manusia dapat menyebabkan iritasi mata dan gangguan pernafasan. Berdasarkan PP 41 tahun 1999 baku mutu konsentrasi ozon yang masih diijinkan adalah tidak lebih dari 235 $\mu\text{g}/\text{m}^3$ untuk waktu pengukuran 1 jam (SARPEDAL KLH, 2003b).

Ozon merupakan salah satu zat pengoksidasi yang sangat kuat setelah fluor, oksigen dan oksigen fluorida (OF_2). Meskipun di alam terdapat dalam jumlah kecil tetapi lapisan lain dengan bahan pencemar udara Ozon sangat berguna untuk melindungi bumi dari radiasi ultraviolet (UV-B). Ozon terbentuk diudara pada ketinggian 30 km. Ozon menyerap radiasi sinar matahari dengan kuat didaerah panjang gelombang 240-320 nm.

2.3 Pencemaran Udara

Udara adalah suatu kesatuan ruangan, dimana makhluk hidup berada didalamnya. Udara atmosfir merupakan campuran gas yang terdiri dari 78% nitrogen, 20% oksigen, 0,93% argon, 0,03% karbon monoksida dan sisanya terdiri dari neon, helium, metan dan hidrogen. Udara dikatakan normal dan dapat mendukung kehidupan manusia, apabila komposisinya seperti tersebut di atas. Pencemaran udara adalah peristiwa masuknya, atau tercampurnya polutan (unsur-unsur berbahaya) ke dalam lapisan udara (atmosfir) yang dapat mengakibatkan menurunnya kualitas udara.

Polutan yang umum mencemari udara berupa gas dan asap. Gas dan asap tersebut berasal dari hasil proses pembakaran bahan bakar yang tidak sempurna, terutama yang dihasilkan oleh mesin-mesin pabrik, pembangkit listrik dan kendaraan bermotor yang tersebar dijalanan. Selain sumber tersebut, gas dan asap juga merupakan hasil oksidasi dari berbagai unsur penyusun bahan bakar, yaitu: CO_2 (karbon dioksida), CO (karbon monoksida), SO_x (belerang oksida) dan NO_x (nitrogen oksida).

2.4 Distribusi Peluang

Distribusi peluang merupakan bagian penting dalam menganalisa peluang dari suatu eksperimen. Seluruh kemungkinan yang terjadi dari titik percobaan dan nilai peluangnya memiliki hubungan yang erat. Secara keseluruhan himpunan

pasangan terurut dari titik percobaan dan nilai peluang padanannya inilah yang disebut sebagai distribusi peluang.

Definisi 2.1 Supranto, 2009 Himpunan pasangan terurut $x, f(x)$ dikatakan sebagai suatu fungsi kepadatan peluang atau distribusi peluang peubah acak diskrit X jika untuk setiap kemungkinan hasil x memenuhi syarat berikut :

1. $f(x) \geq 0$
 2. $\sum f(x) = 1$ (2.1)
 3. Fungsi distribusi kumulatifnya didefinisikan sebagai
- $$F(x) = P(X \leq x)$$

Definisi 2.2 (Harinaldi, 2005) Fungsi $f(x)$ yang kurva distribusi peluangnya diwakili oleh poligon frekuensi relatif yang dimuluskan adalah fungsi kepadatan peluang peubah acak kontinu X , jika memenuhi syarat berikut:

1. $P(a < X < b) = \int_a^b f(x) dx$
2. $f(x) \geq 0$, untuk semua $x \in R$ (non-negatif) (2.2)
3. $\int_{-\infty}^{\infty} f(x) dx = 1$

2.5 Nilai Harapan Distribusi Peluang

Ukuran pemusatan data dari suatu populasi variabel acak disebut sebagai nilai harapan. Nilai harapan ini merupakan ukuran pemusatan data populasi yang sangat penting. Nilai rata-ran peubah acak X atau rata-ran distribusi peluang X ditulis dengan μ_x atau μ . Rataan ini disebut dengan nilai harapan matematik atau nilai harapan peubah acak X dan dinyatakan dengan $E(x)$.

Definisi 2.3 (Harinaldi, 2005) Jika X adalah suatu variabel acak diskrit yang mengambil nilai $x_1, x_2, x_3, \dots, x_n$ yang mempunyai peluang $p_{x_1}, p_{x_2}, p_{x_3}, \dots, p_{x_n}$ dengan $p_{x_1} + p_{x_2} + p_{x_3} + \dots + p_{x_n} = 1$, maka nilai harapan atau rata-ran X untuk variabel acak diskrit adalah :

$$\mu = E(x) = \sum_{i=1}^n x_i P(X = x_i) \quad (2.3)$$

bila variabel acak X kontinu maka nilai harapan X adalah

$$\mu = E X = \int_{-\infty}^{\infty} x f(x) dx, \quad (2.4)$$

Menentukan nilai harapan pada peubah acak diskrit dapat dilakukan dengan mengalikan tiap nilai x_1, x_2, \dots, x_n dari peubah acak X dengan nilai peluangnya $f(x_1), f(x_2), \dots, f(x_n)$, selanjutnya dijumlahkan hasilnya sebagai berikut:

$$p x_1 + p x_2 + p x_3 + \dots + p x_n = 1.$$

Metode ini hampir sama dengan menentukan nilai harapan pada peubah acak kontinu. Nilai harapan dicari dengan menggunakan metode integral untuk mengganti penjumlahan. Penggunaan metode integral ini tidak merubah definisi dari nilai harapan matematik.

2.6 Variansi Distribusi Peluang

Fungsi $g(x) = (X - \mu)^2$ digunakan untuk memperoleh ukuran keragaman atau variansi dari suatu peubah acak X , ukuran keragaman ini selanjutnya biasa disebut sebagai variansi distribusi peluang X yang ditulis dengan $Var x = \sigma_x^2 = \sigma^2$. distribusi peluang X dapat dinyatakan dengan $var(x)$ (Dudewicz & Misra, 1988).

Definisi 2.4 (Dudewicz & Misra, 1988) Diberikan X merupakan peubah acak dengan distribusi peluang $f(x)$ dan nilai harapan μ , maka variansi x untuk variabel acak diskrit X ditentukan oleh

$$Var(x) = E[(x - \mu)^2] = \sum_x x - \mu^2 f(x), \text{ bila } X \text{ diskrit} \quad (2.5)$$

Variansi X pada variabel acak kontinu adalah berikut:

$$Var(x) = E[(x - \mu)^2] = \int_{-\infty}^{\infty} x - \mu^2 f(x) dx, \text{ bila } X \text{ kontinu} \quad (2.6)$$

Teorema 2.1 Variansi dari peubah acak X adalah (Dudewicz & Misra, 1988):

$$Var x = E x^2 - [E(x)]^2 \quad (2.7)$$

2.7 Distribusi Pareto

Distribusi Pareto pertama kali diteliti oleh seorang ahli fisika Italia yang bernama Vilfredo Pareto. Distribusi ini pertama kali digunakan untuk pemodelan

pendapatan dan populasi penduduk kota. Distribusi pareto lebih sering digunakan untuk pemodelan data ekstrim seperti hidrologi, klimatologi, dan udara. Dua parameter yaitu α dan k merupakan ciri dari distribusi pareto (Joseph Lee Petersen).

Distribusi Pareto merupakan distribusi acak kontinu dengan fungsi kepadatan peluang sebagai berikut :

$$f(x, \alpha, k) = \frac{\alpha k^\alpha}{x^{\alpha+1}}; \quad k \leq x < \infty; \quad \alpha, k > 0 \quad (2.8)$$

Berikut akan ditunjukkan apakah fungsi kepadatan peluang distribusi pareto memenuhi syarat suatu fungsi kepadatan peluang, yaitu :

$$1. f(x) \geq 0, \forall x$$

$$f(x, \alpha, k) = \frac{\alpha k^\alpha}{x^{\alpha+1}}; \quad k \leq x < \infty; \quad \alpha, k > 0, \text{ maka terbukti bahwa } f(x) \geq 0, \forall x$$

$$2. \text{ Selanjutnya akan dibuktikan bahwa } \int_{-\infty}^{\infty} f(x) dx = 1$$

Bukti :

$$\begin{aligned} \int_{-\infty}^{\infty} f(x, \alpha, k) dx &= \int_{-\infty}^{\infty} \frac{\alpha k^\alpha}{x^{\alpha+1}} dx \\ &= \int_{-\infty}^k \frac{\alpha k^\alpha}{x^{\alpha+1}} dx + \int_k^{\infty} \frac{\alpha k^\alpha}{x^{\alpha+1}} dx \\ &= 0 + \int_k^{\infty} \frac{\alpha k^\alpha}{x^{\alpha+1}} dx \\ &= \alpha k^\alpha \int_k^{\infty} \frac{1}{x^{\alpha+1}} dx \\ &= \alpha k^\alpha \left[-\frac{1}{\alpha x^\alpha} \right]_k^{\infty} \\ &= 1 \end{aligned}$$

fungsi distribusi kumulatif distribusi pareto adalah :

$$F(x, \alpha, k) = 1 - \frac{k^\alpha}{x^\alpha}; \quad k \leq x < \infty; \quad \alpha, k > 0 \quad (2.9)$$

Selanjutnya akan ditunjukkan fungsi distribusi kumulatif untuk distribusi Pareto sebagai berikut :

$$\begin{aligned} F(t) &= \int_{-\infty}^t f(t) dt \\ &= \int_k^t \frac{\alpha k^\alpha}{t^{\alpha+1}} dt \end{aligned}$$

$$\begin{aligned}
&= ak^a - \frac{t}{a} \frac{k^a}{t^{a+1}} \\
&= - \frac{t k^a}{t^{a+1}} \\
&= - \frac{k^a}{x^a} + 1 = 1 - \frac{k}{x}^a
\end{aligned} \tag{2.10}$$

Selanjutnya akan ditunjukkan rata-rata distribusi Pareto. Rata-rata atau $E(x)$ dari distribusi Pareto adalah sebagai berikut:

$$E(x) = \frac{ak}{(a-1)} \tag{2.11}$$

$$\begin{aligned}
E(x) &= \int_{-\infty}^{\infty} x f(x) dx \\
&= \int_{-\infty}^{\infty} x \frac{ak^a}{x^{a+1}} dx \\
&= \int_{-\infty}^k x \frac{ak^a}{x^{a+1}} dx + \int_k^{\infty} x \frac{ak^a}{x^{a+1}} dx \\
&= 0 + \int_k^{\infty} x \frac{ak^a}{x^{a+1}} dx \\
&= ak^a \int_k^{\infty} \frac{x}{x^{a+1}} dx \\
&= ak^a \int_k^{\infty} x^{-a} dx \\
&= ak^a \left[\frac{1}{-a+1} x^{-a+1} \right]_k^{\infty} \\
&= \frac{ak}{(a-1)}
\end{aligned} \tag{2.12}$$

Selanjutnya akan ditunjukkan variansi distribusi Pareto yaitu :

$$\begin{aligned}
Var(x) &= E(x^2) - [E(x)]^2 \\
Var(x) &= \frac{ak^2}{a-1} \frac{1}{(a-2)}
\end{aligned} \tag{2.13}$$

Sebelum menggunakan Persamaan (2.7), akan ditentukan lebih dulu $E(x^2)$ sebagai berikut :

$$\begin{aligned}
E(x^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx \\
&= \int_{-\infty}^{\infty} x^2 \frac{ak^a}{x^{a+1}} dx, \text{ bentuk integral ini dapat dipecah menjadi :} \\
&= \int_{-\infty}^k x^2 \frac{ak^a}{x^{a+1}} dx + \int_k^{\infty} x^2 \frac{ak^a}{x^{a+1}} dx \\
&= 0 + \int_k^{\infty} x^2 \frac{ak^a}{x^{a+1}} dx
\end{aligned}$$

$$\begin{aligned}
&= \alpha k^\alpha \int_k^\infty \frac{x^2}{x^{\alpha+1}} dx \\
&= \alpha k^\alpha \int_k^\infty \frac{1}{x^{\alpha+2}} dx \\
&= \frac{\alpha k^2}{\alpha-2}
\end{aligned}$$

Substitusikan Persamaan $E x^2$ dan $[E(x)]^2$ yang diperoleh diatas, maka diperoleh variansi distribusi Pareto, yaitu :

$$\begin{aligned}
Var x &= E x^2 - [E(x)]^2 \\
&= \frac{\alpha k^2}{\alpha-2} - \left(\frac{\alpha k}{\alpha-1} \right)^2 \\
&= \frac{\alpha k^2 \alpha^2 k^2}{(\alpha-2)(\alpha-1)} \\
&= \frac{\alpha k^2}{(\alpha-2)^2 (\alpha-1)}
\end{aligned} \tag{2.14}$$

2.8 Distribusi Nilai Ekstrim

Fenomena alam dan gejala alam merupakan hal yang sangat menarik untuk dikaji. Dua hal ini banyak dikaji mulai dari fenomena dan gejala alam yang biasa sampai yang bersifat ekstrim. Fenomena alam yang banyak dibicarakan ini meliputi angin, banjir, hujan, cuaca dan iklim serta yang berhubungan dengan udara seperti ozon. Kejadian ekstrim yang sekarang ini sedang banyak dibicarakan adalah kondisi ozon yang berpotensi memicu pemanasan global. Kondisi ozon yang ekstrim dalam jangka waktu yang lama dan melampaui batas aman udara bersih akan sangat berpengaruh pada kehidupan di bumi.

Ahli meteorologi sekarang ini telah dilengkapi dengan radar, sinar laser dan komputer serta kecanggihan teknologi yang mampu memprediksi kondisi kejadian ekstrim dimasa yang akan datang dengan baik. Demikian juga dengan ozon ekstrim, kondisi ini juga bisa diperbaiki dimasa yang akan datang dengan mengurangi tingkat pencemaran udara. Dalam mengantisipasi akibat buruk yang akan terjadi mendatang, dibutuhkan informasi dan model yang bisa digunakan untuk simulasi data dan memprediksi kejadian-kejadian ekstrim yang akan terjadi.

Dengan demikian akan bisa diambil tindakan yang dapat mencegah dampak buruk yang akan terjadi.

Adanya model distribusi yang dapat menggambarkan kejadian ekstrim pada data pencemaran udara oleh ozon akan sangat membantu keakuratan pendugaan parameter dan simulasi untuk data ekstrim berikutnya. Salah satu distribusi nilai ekstrim yang baik digunakan adalah distribusi *Generalized Extreme Value* (GEV). GEV merupakan suatu distribusi yang dikembangkan untuk mengkaji kejadian-kejadian ekstrim. Oleh karena itu distribusi GEV ini sangat bermanfaat untuk memodelkan data ekstrim (Jantje Denny Prang, 2006).

Enrick Castillo dalam penelitiannya menyatakan bahwa distribusi *Generalized extreme value (GEV)* termasuk kedalam distribusi acak kontinu dan diterbitkan dalam buku "*Extreme Value and Related Models with Applications in Engineering and Science*". Dalam bukunya Enrick Castillo menyatakan distribusi GEV baik untuk memodelkan data ekstrim (Enrick Castillo, 2009).

Penelitian terkait lainnya adalah "*Mixed Methods for Fitting the GEV Distribution*" yang ditulis oleh Pierre dan Craig pada Tahun 2008. Dalam penelitiannya disebutkan bahwa distribusi GEV banyak digunakan untuk pomodelan data yang bersifat ekstrim. Hal ini dikarenakan distribusi ini mempunyai tiga parameter yang merupakan gabungan dari tiga distribusi yaitu Pareto, Freceth, dan Weibull. Distribusi ini termasuk kedalam distribusi kontinu yang dikhususkan untuk memodelkan data ekstrim (Pierre dan Craig, 2008).

Model statistik untuk data ekstrim maksimum adalah teori nilai ekstrim yang disebut *Generalized Extreme Value (GEV)*. Pada umumnya distribusi ini memang digunakan untuk memodelkan data ekstrim. Data ekstrim yang dipakai berada dalam rentang maksimum jangka waktu tertentu seperti dalam skala harian, bulanan, dan tahunan. Pada kenyataanya data ekstrim maksimum memang sangat berguna untuk dijadikan acuan dalam tindakan pencegahan kejadian ekstrim mendatang (Gilli dan Kellezi, 2006).

Pembuktian syarat kontinu distribusi GEV menggunakan cara yang sama pada pembuktian syarat kontinu pada distribusi pareto pada persamaan (2.10),

(2.12), (2.13). Sehingga dari fungsi kepadatan peluang distribusi *generalized extreme value (GEV)* berikut:

$$f(x, \mu, \sigma, \xi) = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{\frac{1}{\xi-1}} e^{- \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{\frac{1}{\xi}}}, \quad \xi \neq 0 \quad (2.15)$$

$$f(x, \mu, \sigma, \xi) = \frac{1}{\sigma} e^{\frac{x - \mu}{\sigma}} e^{- e^{\frac{x - \mu}{\sigma}}}, \quad \xi = 0$$

Diperoleh fungsi distribusi kumulatifnya yaitu,

$$F(x) = \begin{cases} e^{- \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{\frac{1}{\xi}}}, & \xi \neq 0 \\ e^{- e^{\frac{x - \mu}{\sigma}}}, & \xi = 0 \end{cases} \quad (2.16)$$

Sehingga berdasarkan penelitian dari Jantje Denny Prang, Enrick Castillo, Pierre dan Craig, serta Gilli dan Kellezi maka jelas bahwa distribusi GEV merupakan distribusi yang termasuk kedalam distribusi acak kontinu dan merupakan distribusi yang baik dan cocok untuk memodelkan data ekstrim.

2.9 Estimasi Parameter

Dalam menentukan model distribusi yang sesuai untuk suatu data, terlebih dahulu ditentukan parameter dari distribusi tersebut. Metode yang digunakan salah satunya adalah metode maksimum *likelihood*. Metode maksimum *likelihood* sering digunakan dalam penelitian karena prosedur atau langkah-langkahnya sangat jelas dan sesuai dalam menentukan parameter dari sebuah distribusi (Krishnamoorthy, 2006).

2.10 Fungsi Likelihood

Fungsi likelihood dapat dijelaskan dengan fungsi kepadatan peluang yang dievaluasi pada setiap titik x_1, x_2, \dots, x_n . Fungsi kepadatan peluang (FKP) bersama dari variabel acak x_1, x_2, \dots, x_n yaitu $f(x_1, x_2, \dots, x_n; \theta)$ yang dievaluasi pada titik x_1, x_2, \dots, x_n yang disebut fungsi *likelihood* yang dinotasikan dengan $L(\theta; X)$ maka :

$$L(\theta; X) = f(X; \theta) \quad (2.17)$$

karena $f(x_1, x_2, \dots, x_n; \theta)$ adalah FKP bersama dari variabel acak yang saling bebas, sehingga :

$$f(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta) \quad (2.18)$$

Selanjutnya substitusikan Persamaan (2.17) ke Persamaan (2.18), maka :

$$\begin{aligned} L(\theta; X) &= f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta) \\ &= \prod_{i=1}^n f(x_i; \theta) \end{aligned} \quad (2.19)$$

2.10.1 Estimasi Maksimum *Likelihood*

Estimasi Maksimum *Likelihood* (EML) adalah suatu metode yang memaksimumkan fungsi *likelihood*. Prinsip yang digunakan untuk estimasi maksimum *likelihood* adalah memilih $\hat{\theta}$ sebagai estimator titik untuk θ yang memaksimumkan $L(\theta; X)$. Metode EML dapat digunakan jika fungsi kepadatan peluang (FKP) telah diperoleh.

Misalkan x_1, x_2, \dots, x_n adalah sampel acak dari suatu distribusi dengan FKP $f(X; \theta)$, kemudian dibentuk FKP bersama x_1, x_2, \dots, x_n , setelah itu ditentukan fungsi *likelihood* dari θ yaitu $L(\theta; X)$. Metode estimasi maksimum *likelihood* membuat fungsi *likelihood* $L(\theta; X)$ menjadi maksimum dan digunakan fungsi logaritma. Sehingga fungsi logaritma *likelihood* dinotasikan dengan $\ln L(\theta; X) = l(\theta; X)$, dimana $l(\hat{\theta}; X) = l(\theta; X)$. Dengan menggunakan logaritma $L(\theta; X)$, maka estimator *likelihood* diperoleh dari turunan fungsi *likelihood* terhadap parameternya, yaitu $\frac{dl(\theta; X)}{d\theta} = 0$.

2.10.2 Estimasi Parameter Menggunakan Metode Maksimum *Likelihood*

Metode maksimum *likelihood* adalah salah satu metode yang digunakan dalam menentukan parameter dari sebuah distribusi. Dalam penelitian ini akan digunakan metode tersebut untuk menentukan parameter dari distribusi Pareto.

Parameter dari fungsi kepadatan peluang Pareto (α, k) pada Persamaan (2.8) dapat ditunjukkan sebagai berikut :

$$f(x, \alpha, k) = \frac{\alpha k^\alpha}{x_i^{\alpha+1}}; \quad k < x < \infty; \quad \alpha, k > 0$$

maka fungsi *likelihood* dari Persamaan (2.8) sebagai berikut :

$$\begin{aligned} L &= f(x_1) f(x_2) \dots f(x_n) \\ &= \frac{\alpha k^\alpha}{x_1^{\alpha+1}} \frac{\alpha k^\alpha}{x_2^{\alpha+1}} \dots \frac{\alpha k^\alpha}{x_n^{\alpha+1}} \\ &= \frac{\alpha^n k^{\alpha n}}{\prod_{i=1}^n x_i^{\alpha+1}} \end{aligned} \quad (2.20)$$

Setelah diperoleh fungsi *likelihood*, langkah berikutnya akan ditentukan maksimum *likelihood* dari Persamaan (4.1) dengan menjadikan fungsi *likelihood* tersebut menjadi logaritma *likelihood*, yaitu :

$$\begin{aligned} l &= \ln L \\ &= \ln \frac{\alpha^n k^{\alpha n}}{\prod_{i=1}^n x_i^{\alpha+1}} \\ l &= \ln \alpha^n + \ln k^{\alpha n} - \sum_{i=1}^n \ln \frac{1}{x_i^{\alpha+1}} \\ &= n \ln \alpha + \alpha n \ln k - \sum_{i=1}^n (\alpha + 1) \ln(x_i) \end{aligned} \quad (2.21)$$

karena nilai dari $\frac{\partial(\alpha, X)}{\partial \alpha} = 0$,

sehingga diperoleh,

$$\begin{aligned} \frac{\partial l}{\partial \alpha} &= \frac{n}{\alpha} + n \ln k - \sum_{i=1}^n \ln(x_i) = 0 \\ \frac{n}{\alpha} &= -n \ln k + \sum_{i=1}^n \ln(x_i) \end{aligned}$$

$$\begin{aligned}
\alpha &= \frac{n}{-n \ln k + \sum_{i=1}^n \ln(x_i)} \\
\alpha &= \frac{n}{-(\ln(k^n) + \sum_{i=1}^n \ln(x_i))} \\
\alpha &= \frac{n}{-\ln \sum_{i=1}^n k^n + \sum_{i=1}^n \ln(x_i)} \\
\alpha &= \frac{n}{\sum_{i=1}^n \ln(x_i) - \sum_{i=1}^n \ln(k^n)} \\
\alpha &= \frac{n}{\sum_{i=1}^n (\ln(x_i) - \ln(k))} \\
\alpha &= \frac{n}{\sum_{i=1}^n \ln\left(\frac{x_i}{k}\right)} \tag{2.22}
\end{aligned}$$

Parameter k dalam distribusi Pareto tidak perlu diturunkan dari fungsi *likelihood* dengan cara yang sama, karena parameter k merupakan nilai yang terkecil dari x_i pada data penelitian. Oleh karena itu nilai parameter k dapat diperoleh langsung dari data. Sehingga parameter k dapat ditulis sebagai berikut :

$$k = \min x_i \tag{2.23}$$

dengan x_i adalah data ozon ekstrim di Pekanbaru Tahun 2009-2012

2.10.3 Estimasi *Probability Weighted Moments (PWM)*

Fungsi PWM dari variabel random x dengan CDF $F(x)$ adalah

$$M_{p,r,s} = E[x^p F(x)^r (1 - F(x))^s] \tag{2.24}$$

Dengan $p, r, s =$ bilangan real. Langkah-langkah estimasi dengan *probability weighted moments* adalah:

- 1) Memformulasikan fungsi PWM β_r dengan $r = 0,1,2$
- 2) Memformulasikan estimator *unbiased* untuk fungsi PWM $\hat{\beta}_r$ dengan $r = 0,1,2$
- 3) Menghitung $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ dari fungsi PWM
- 4) Hasil persamaan yang diperoleh dari $\hat{\beta}_0$ digunakan untuk memperoleh μ

- 5) Menghitung $2\beta_1 - \beta_0$ dan $3\beta_2 - \beta_0$ dari fungsi PWM sehingga diperoleh $2\bar{\beta}_1 - \bar{\beta}_0$ dan $3\bar{\beta}_2 - \bar{\beta}_0$
- 6) Hasil yang diperoleh dari $2\bar{\beta}_1 - \bar{\beta}_0$ digunakan untuk mendapatkan nilai $\hat{\sigma}$
- 7) Membuat perbandingan $3\bar{\beta}_2 - \bar{\beta}_0$ dan $2\bar{\beta}_1 - \bar{\beta}_0$
- 8) Hasil persamaan yang diperoleh melalui perbandingan $3\bar{\beta}_2 - \bar{\beta}_0$ dan $2\bar{\beta}_1 - \bar{\beta}_0$ (Anita Rahayu, 2012).

2.10.4 Estimasi Parameter dengan Menggunakan *Probability Weight Moment (PWM)*

Fungsi PWM dari variabel random X dengan *Cumulative Distribution Function* $F(x)$ adalah :

$$M_{p,r,s} = E [X^p F(X)^r (1 - F(X))^s] \text{ dengan } p, r, s = \text{bilangan real} \quad (2.25)$$

$$\begin{aligned} \beta_r &= M_{1,r,0} \\ &= E [F(X)^r] \\ &= \frac{1}{r+1} \left[\mu + \frac{\sigma}{\xi} \left(1 - \frac{r}{r+1} \right)^{-\xi} - 1 \right] \quad , \xi < 1, \xi \neq 0 \end{aligned} \quad (2.26)$$

Selanjutnya akan ditentukan estimator unbiased dari β_r sebagai berikut:

$$\begin{aligned} \hat{\beta}_r &= \bar{M}_{1,r,0} \\ &= \frac{1}{n} \sum_{j=1}^n \sum_{l=1}^n \frac{j-l}{n-l} x_j \\ &= \frac{1}{n} \sum_{j=1}^n \frac{j-1}{n-1} \frac{j-2}{n-2} \dots \frac{j-r}{n-r} x_j \end{aligned} \quad (2.27)$$

$$\hat{\beta}_0 = \bar{x} \quad (2.28)$$

$$\hat{\beta}_1 = \frac{1}{n} \sum_{j=1}^n \frac{j-1}{n-1} x_j \quad (2.29)$$

$$\hat{\beta}_2 = \frac{1}{n} \sum_{j=1}^n \frac{j-1}{n-1} \frac{j-2}{n-2} x_j \quad (2.30)$$

$$\hat{\beta}_0 = \frac{1}{0+1} \left(\mu + \frac{\sigma}{\xi} \left(1 - \frac{0+1}{1+\xi} \right) \right)$$

$$\hat{\beta}_0 = \mu + \frac{\sigma}{\xi} \left(1 - \frac{1}{1+\xi} \right)$$

$$\hat{\beta}_0 = \hat{\mu} + \frac{\hat{\sigma}}{\xi} \left(1 - \frac{1}{1+\xi} \right)$$

$$\hat{\mu} = \hat{\beta}_0 - \frac{\hat{\sigma}}{\xi} \left(1 - \frac{1}{1+\xi} \right) \quad (2.31)$$

selanjutnya akan ditentukan β_1 dan β_2 , setelah itu akan dibandingkan untuk memperoleh parameter ξ .

$$\beta_1 = \frac{1}{2} \left(\mu + \frac{\sigma}{\xi} \left(1 - \frac{2}{1+\xi} \right) \right)$$

$$2\beta_1 - \beta_0 = \frac{\sigma}{\xi} \left(1 + \frac{\xi}{1+\xi} \right) \left(1 - \frac{2}{1+\xi} \right)$$

$$\beta_2 = \frac{1}{3} \left(\mu + \frac{\sigma}{\xi} \left(1 - \frac{3}{1+\xi} \right) \right)$$

$$3\beta_2 - \beta_0 = \frac{\sigma}{\xi} \left(1 + \frac{\xi}{1+\xi} \right) \left(1 - \frac{3}{1+\xi} \right)$$

$$\frac{3\beta_2 - \beta_0}{2\beta_1 - \beta_0} = \frac{1 - \frac{3}{1+\xi}}{1 - \frac{2}{1+\xi}}$$

$$\xi = 7,8590 c + 2,9554 c^2 \quad (2.32)$$

$$\text{dengan } c = \frac{2\hat{\beta}_1 - \hat{\beta}_0}{3\hat{\beta}_2 - \hat{\beta}_0} - \frac{\ln 2}{\ln 3} \quad (2.33)$$

Nilai c diperoleh berdasarkan penelitian dari Samuel Kotz. Berikutnya akan ditentukan anggaran parameter untuk $\hat{\sigma}$, sebagai berikut :

$$2\hat{\beta}_1 - \hat{\beta}_0 = \frac{\hat{\sigma}}{\xi} \left(1 + \frac{\xi}{1+\xi} \right) \left(1 - \frac{2}{1+\xi} \right)$$

$$2\hat{\beta}_1 - \hat{\beta}_0 \xi = \hat{\sigma} \left(1 + \frac{\xi}{1+\xi} \right) \left(1 - \frac{2}{1+\xi} \right)$$

$$\hat{\sigma} = \frac{2\hat{\beta}_1 - \hat{\beta}_0}{1 + \frac{\xi}{1 - 2^{-\xi}}} \quad (2.34)$$

2.11 Uji Kebaikan (*Goodness of Fit*)

Uji kebaikan (*Goodness of Fit*) adalah uji yang dilakukan untuk memperoleh model distribusi yang sesuai terhadap data observasi yang digunakan dalam sebuah penelitian. Uji kebaikan digunakan berdasarkan fungsi distribusi kumulatif secara lengkap dengan parameter-parameter yang telah ditentukan. Pada penelitian ini, model distribusi yang sesuai untuk data akan ditentukan dengan menggunakan uji Kolmogorov-Smirnov dan Anderson-Darling A^2 (Thode, 2002).

2.11.1 Uji Kolmogorov-Smirnov

Statistik ini menggunakan fungsi distribusi kumulatif dan berdasarkan pada perbedaan maksimum antara dua distribusi, yaitu distribusi normal dengan distribusi data yang diamati. Uji statistik Kolmogorov-Smirnov ditunjukkan pada persamaan berikut :

$$D = \max D^+, D^- \quad (2.35)$$

dimana,

$$D^+ = \max_{i=1, \dots, n} \frac{i}{n} - F(x_i) \quad (2.36)$$

dan,

$$D^- = \max_{i=1, \dots, n} F(x_i) - \frac{(i-1)}{n} \quad (2.37)$$

dengan $F(x_i)$ adalah fungsi distribusi kumulatif. Nilai D berdasarkan pada jarak maksimum antara D^+ dan D^- . Model distribusi dikatakan sesuai untuk data jika uji statistik D pada suatu model distribusi tersebut bernilai minimum (Thode, 2002).

2.11.2 Uji Anderson-Darling

Uji statistik ini pertama kali dikembangkan oleh Anderson dan Darling pada tahun 1954. Uji statistik Anderson–Darling berdasarkan pada fungsi distribusi empirik. Biasanya digunakan untuk data berukuran besar.

Uji Anderson-Darling ditunjukkan pada persamaan berikut :

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \frac{2i-1}{2n+1-2i} \ln \frac{F(x_i)}{1-F(x_i)} \quad (2.38)$$

dengan $F(x_i)$ adalah fungsi distribusi kumulatif. Model distribusi dikatakan sesuai untuk data jika uji statistik A^2 pada suatu model distribusi tersebut bernilai minimum (Pani, 2009).

2.12 Penelitian-penelitian yang Terkait dengan Model Distribusi

Penelitian-penelitian yang berhubungan dengan model distribusi yang pernah diteliti sebelumnya, antara lain :

Tabel 2.1 Penelitian-Penelitian Model Distribusi

No	Judul Penelitian	Peneliti	Metode	Tahun
1.	<i>Model Statistika untuk Data Karbon Monoksida (CO)</i>	Ari Pani Desvina	<i>Exponensial, Weibull, Lognormal, Distribusi Gamma, Distribusi Nilai ekstrem Teritlak</i>	2009
2.	<i>Statistical Analysis of Extreme Ozone Data</i>	Norhaslinda binti Ali, dkk	<i>Distribusi nilai ekstrem, poisson</i>	2012
3.	<i>Pemeriksaan Data Berpengaruh Dalam Model Regresi Gamma</i>	Nusar Hajarisman	<i>Distribusi gamma</i>	2010

BAB III

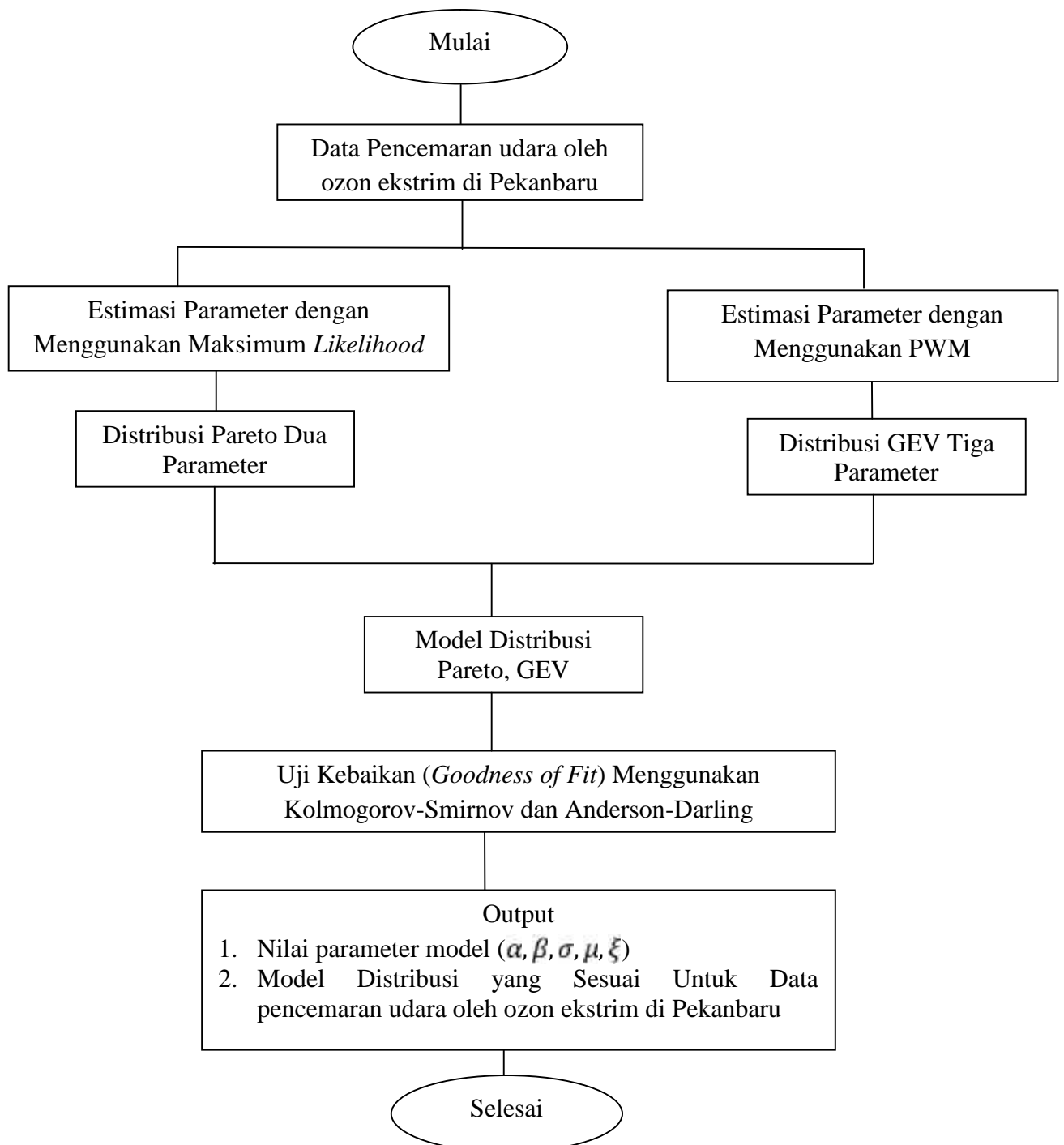
METODOLOGI PENELITIAN

Metode yang digunakan dalam penulisan tugas akhir ini adalah studi pustaka. Metode ini menggunakan literatur yang berkaitan dengan permasalahan yang dibahas. Bab ini juga berisikan tentang informasi data yang digunakan dan sumber data.

Langkah-langkah dalam penyelesaian tugas akhir ini dilakukan sebagai berikut:

1. Mengumpulkan data pencemaran udara ozon ekstrim dalam skala harian dari Tahun 2009-2012 di Pekanbaru yang diperoleh dari Badan Lingkungan Hidup Kota Pekanbaru.
2. Data yang telah dikumpulkan selanjutnya diorganisir hingga data siap untuk dianalisis dengan menggunakan *software* statistik yaitu *easyfit*.
3. Menentukan nilai parameter distribusi pareto dan distribusi GEV dengan menggunakan *software easyfit*.
4. Membangun model statistik berdasarkan nilai parameter yang telah diperoleh pada langkah tiga untuk distribusi Pareto dan distribusi GEV.
5. Melakukan uji *goodness of fit* untuk distribusi Pareto dan GEV dengan menggunakan uji Kolmogorov Smirnov dan Anderson Darling untuk menentukan distribusi yang terbaik.
6. Menentukan model terbaik berdasarkan uji *goodness of fit* yang digunakan.
7. Selesai.

Langkah-langkah pengumpulan data dan membangun model diatas dapat digambarkan dalam *flow chart* sebagai berikut :



Gambar 3.1 *Flowchart* Metodologi Penelitian

BAB IV

PEMBAHASAN

4.1 Deskriptif Statistik Data

Data yang digunakan dalam penelitian ini adalah data pencemaran udara oleh ozon ekstrim di Pekanbaru dari Tahun 2009-2012. Data diambil dalam skala maksimum harian. Data yang nilainya paling maksimum dalam skala harian dinyatakan sebagai data ekstrim. Keseluruhan data dalam penelitian ini berjumlah 1362 data yang telah diurutkan dari yang terkecil sampai yang terbesar. Bentuk data kepekatan ozon ekstrim di Pekanbaru dari Tahun 2009-2012 diberikan dalam tabel berikut:

Tabel 4.1 Data Pencemaran Udara oleh Ozon Ekstrim

No	Hari/Tanggal	Rata-rata Harian
1	Kamis, 1 Januari 2009	3,9
2	Jumat,2 Januari 2009	4,8
3	Sabtu,3 Januari 2009	4,97
4	Minggu,4 Januari 2009	5,22
5	Senin,5 Januari 2009	6,19
6	Selasa,6 Januari 2009	7,48
7	Rabu,7 Januari 2009	7,49
8	Kamis,8 Januari 2009	7,59
9	Jumat,9 Januari 2009	7,68
10	Sabtu,10 Januari 2009	8,3
11	Minggu,11 Januari 2009	9,06
12	Senin,12 Januari 2009	9,22
13	Selasa,13 Januari 2009	9,49
14	Rabu,14 Januari 2009	9,77
15	Kamis,15 Januari 2009	9,88
16	Jumat,16 Januari 2009	10,02
17	Sabtu,17 Januari 2009	10,05
⋮	⋮	⋮
1362	Senin,31 Desember 2012	128,92

Sumber: Badan Lingkungan Hidup Kota Pekanbaru

4.2 Menentukan Nilai Parameter Data Ozon Ekstrim

Setelah diperoleh persamaan parameter dari distribusi Pareto dan distribusi GEV. Selanjutnya akan ditentukan nilai parameter tersebut dari data ozon ekstrim di Pekanbaru.

a. Distribusi Pareto

Nilai parameter α pada distribusi Pareto dapat dicari dengan Persamaan yang telah diperoleh pada Persamaan (2.22), yaitu :

$$\alpha = \frac{n}{\sum_{i=1}^n \ln \frac{x_i}{k}}$$

$$\alpha = \frac{1362}{3011,055548}$$

$$\alpha = 0,45233$$

selanjutnya nilai parameter k , karena merupakan nilai terkecil dari data maka k dapat dicari dengan menggunakan Persamaan (2.23), yaitu :

$$k = \min x_i$$

$$k = 3,90$$

b. Generalized Extreme Value (GEV)

Nilai parameter distribusi GEV dapat dicari dengan rumus parameter pada Persamaan (2.31), (2.32), (2.34). sebelum mencari nilai parameter μ dan σ terlebih dahulu ditentukan nilai dari parameter ξ yaitu:

$$\xi = 7,8590 c + 2,9554 c^2$$

$$\text{dengan } c = \frac{2\hat{\beta}_1 - \hat{\beta}_0}{3\hat{\beta}_2 - \hat{\beta}_0} - \frac{\ln 2}{\ln 3}$$

Maka harus ditentukan terlebih dahulu nilai dari $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$. Nilai $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ ditentukan dengan menggunakan data pencemaran udara ozon ekstrim Pekanbaru dari tahun 2009-2012. Berdasarkan tabel 4.1 diketahui data pencemaran udara yaitu:

$$x_1 = 3,90$$

$$x_1 = 4,80$$

$$x_1 = 4,97$$

⋮

$$x_n = 128,92$$

Dengan jumlah data $n = 1362$, maka nilai $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ dapat ditentukan sebagai berikut:

$$\begin{aligned}\hat{\beta}_0 &= \bar{x} \\ &= \frac{\sum_{i=1}^n x_i}{n} \\ &= \frac{3,90 + 4,80 + 4,97 + \dots + 128,92}{1362} \\ &= 43,0553\end{aligned}$$

$$\begin{aligned}\hat{\beta}_1 &= \frac{1}{n} \sum_{j=1}^n \frac{j-1}{n-1} x_j \\ &= \frac{1}{1362} \frac{1-1}{1362} 3,90 + \frac{2-1}{1362} 4,80 + \dots + \frac{1362-1}{1362} 128,92 \\ &= 28,6975\end{aligned}$$

$$\begin{aligned}\hat{\beta}_2 &= \frac{1}{n} \sum_{j=1}^n \frac{j-1}{n-1} \frac{j-2}{n-2} x_j \\ &= \frac{1}{1362} \frac{1-1}{1362} \frac{(1-2)}{1362} 3,90 + \frac{2-1}{1362} \frac{(2-2)}{1362} 4,80 + \dots \\ &\quad + \frac{1362-1}{1362} \frac{(1362-2)}{1362} 128,92 \\ &= 22,0299\end{aligned}$$

Sehingga nilai \hat{c} dapat diperoleh sebagai berikut:

$$\begin{aligned}\hat{c} &= \frac{2\hat{\beta}_1 - \hat{\beta}_0}{3\hat{\beta}_2 - \hat{\beta}_0} - \frac{\ln 2}{\ln 3} \\ \hat{c} &= \frac{2 \cdot 28,6975 - 43,0553}{3(22,0299) - 43,0553} - \frac{0,69324}{1,09861} \\ \hat{c} &= -0,001839\end{aligned}$$

Selanjutnya nilai parameter ξ diperoleh sebagai berikut:

$$\hat{\xi} = 7,8590 - 0,001839 + 2,9554 - 0,001839^2$$

$$\hat{\xi} = 0,06755$$

Setelah mendapatkan nilai parameter ξ , akan ditentukan nilai parameter σ dan μ sebagai berikut:

$$\hat{\sigma} = \frac{2\hat{\beta}_1 - \hat{\beta}_0 \hat{\xi}}{\Gamma(1 + \hat{\xi}) - 1 - 2^{-\hat{\xi}}}$$

$$\hat{\sigma} = \frac{2 \cdot 28,6975 - 43,0553 \cdot 0,06755}{\Gamma(1 + 0,06755) - 1 - 2^{-0,06755}}$$

$$\hat{\sigma} = 19,473$$

dan nilai μ diperoleh sebagai berikut:

$$\hat{\mu} = \hat{\beta}_0 - \frac{\hat{\sigma}}{\hat{\xi}} (1 - \Gamma(1 + \hat{\xi}))$$

$$\hat{\mu} = 43,0553 - \frac{19,473}{0,06755} (1 - \Gamma(1 + 0,06755))$$

$$\hat{\mu} = 30,427$$

Nilai parameter dari distribusi Pareto dan GEV dapat dilihat pada tabel berikut:

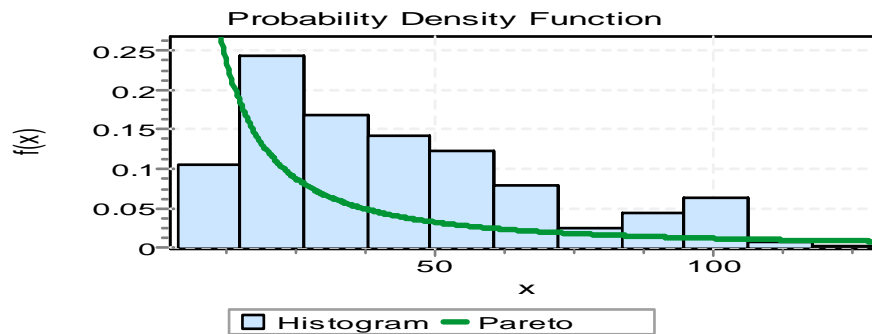
Tabel 4.2 Nilai Parameter Distribusi Pareto dan GEV

No	Distribusi	Parameter				
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$
1	Pareto	0,45233	3,90			
2	GEV			30,427	19,473	0,6755

4.3 Model Distribusi untuk Data Pencemaran Udara oleh Ozon Ekstrim di Pekanbaru

a. Distribusi Pareto

Model distribusi Pareto untuk data pencemaran udara oleh ozon ekstrim di Pekanbaru dari tahun 2009-2012 dengan fungsi kumulatif yang diketahui seperti pada Persamaan dapat digambarkan dengan kurva sebagai berikut :

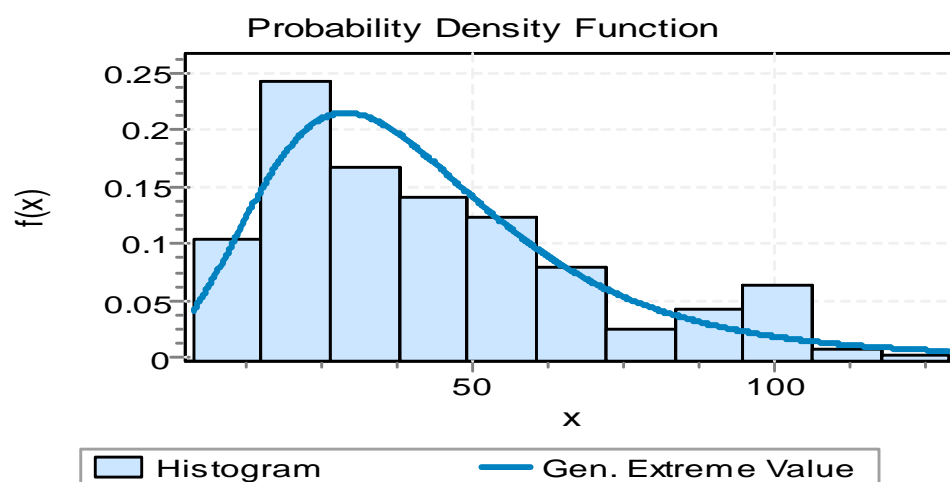


Gambar 4.1 Model Statistik Distribusi Pareto untuk Data Pencemaran Udara oleh Ozon Ekstrim di Pekanbaru

Kurva data pencemaran udara oleh ozon ekstrim dengan distribusi Pareto tidak membentuk kurva yang mendekati kurva normal. Hal ini berarti distribusi Pareto tidak sesuai untuk memodelkan data pencemaran udara oleh ozon ekstrim di Pekanbaru. Karena penelitian ini menggunakan dua distribusi yaitu distribusi Pareto dan GEV, maka perlu diterapkan pula distribusi GEV pada data yang sama.

b. Distribusi *Generalized Extreme Value* (GEV)

Berikut model statistik yang dihasilkan dari distribusi GEV dalam bentuk kurva.



Gambar 4.2 Model Statistik Distribusi GEV untuk Data Pencemaran Udara oleh Ozon Ekstrim di Pekanbaru

Sebagaimana distribusi Pareto, model distribusi GEV dari data pencemaran udara oleh ozon ekstrim di Pekanbaru diperoleh kurva dengan

kelengkungan mendekati kurva normal dan memiliki kemiringan yang lebih mendekati dibandingkan dengan distribusi pareto. Secara grafik distribusi GEV lebih baik dibandingkan dengan distribusi Pareto. Grafik kedua distribusi yang telah diperoleh sudah menunjukkan perbandingan dan distribusi GEV jelas lebih baik dibandingkan dengan Pareto. Kesesuaian distribusi yang diperoleh dapat ditunjukkan keakuratannya dengan uji kebaikan atau *goodness of fit* yaitu menggunakan Kolmogorov Smirnov dan Anderson Darling.

4.4 Uji Kebaikan Model (*Goodness of Fit*)

Kegunaan uji kebaikan model digunakan untuk menunjukkan kesesuaian model yang diperoleh secara matematis. Dengan demikian model yang digunakan benar-benar model yang sesuai.

4.4.1 Uji Kolmogorov Smirnov

a. Distribusi Pareto

Uji Kolmogorov-Smirnov untuk distribusi Pareto dilakukan dengan menggunakan Persamaan (2.24), yaitu :

$$D = \max D^+, D^-$$

$$D^+ = \max_{i=1, \dots, n} \frac{i}{n} - F(x_i)$$

$$D^- = \max_{i=1, \dots, n} \frac{i}{n} - F(x_i)$$

$$D^+ = \max_{i=1, \dots, n} \frac{i}{n} - 1 - \frac{k}{x_i}^\alpha$$

dengan,

$$i = 1, 2, \dots, n$$

$$n = 1362$$

$$x_i = \text{data ke } i$$

$$\alpha = 0,45233$$

$$k = 3,90$$

Sehingga Persamaan menjadi,

$$D^+ = \max_{i=1, \dots, n} \frac{i}{1362} - 1 - \frac{3,90}{x_i}^{0,45233}$$

$$D^+ = 0,21652$$

Selanjutnya, akan ditentukan nilai D^- , sebagai berikut :

$$D^- = \max_{i=1, \dots, n} F(x_i) - \frac{(i-1)}{n}$$

$$D^- = \max_{i=1, \dots, n} 1 - \frac{\frac{k}{x_i}^{\alpha}}{n} - \frac{(i-1)}{n}$$

$$D^- = \max_{i=1, \dots, n} 1 - \frac{3,90}{x_i}^{0,45233} - \frac{(i-1)}{1362}$$

$$D^- = 0,37155$$

maka diperoleh :

$$D = \max D^+, D^-$$

$$D = \max 0,21652, 0,37155$$

$$D = 0,37155$$

b. Distribusi GEV

Berikutnya akan dilakukan uji Anderson Darling untuk distribusi Pareto dengan menggunakan persamaan yaitu:

$$D = \max D^+, D^-$$

$$D^+ = \max_{i=1, \dots, n} \frac{i}{n} - F(x_i)$$

$$D^+ = \max_{i=1, \dots, n} \frac{i}{n} - F(x_i)$$

$$D^+ = \max_{i=1, \dots, n} \frac{i}{n} - e^{-1+\xi \frac{x-\mu}{\sigma}}^{\frac{1}{\xi}}$$

dengan,

$$i = 1, 2, 3 \dots, n$$

$$n = 1362$$

$$x_i = \text{data ke } i$$

$$\mu = 30,427$$

$$\sigma = 19,473$$

$$\xi = 0,06755$$

Substitusikan nilai parameter yang telah diperoleh untuk menyelesaikan uji yang dilakukan. Agar lebih mudah, perhitungan selengkapnya dapat dilihat pada lampiran. Sehingga diperoleh hasil perhitungan sebagai berikut:

$$D^+ = \max_{i=1, \dots, n} \frac{i}{n} - e^{-1 + 0,06755 \frac{x_i - 30,427}{19,473} \frac{1}{0,06755}}$$

$$D^+ = 0,999266$$

Selanjutnya dicari nilai dari $D^- = \max_{i=1, \dots, n} F(x_i) - \frac{(i-1)}{n}$ sebagai berikut:

$$D^- = \max_{i=1, \dots, n} F(x_i) - \frac{(i-1)}{n}$$

$$D^- = \max_{i=1, \dots, n} e^{-1 + \xi \frac{x_i - \mu}{\sigma} \frac{1}{\xi}} - \frac{(i-1)}{n}$$

dengan,

$$i = 1, 2, 3 \dots, n$$

$$n = 1363$$

$$x_i = \text{data ke } i$$

$$\mu = 30,427$$

$$\sigma = 19,473$$

$$\xi = 0,06755$$

Sehingga Persamaan menjadi sebagai berikut:

$$D^- = \max_{i=1, \dots, n} e^{-1 + 0,06755 \frac{x_i - 30,427}{19,473} \frac{1}{0,06755}} - \frac{(i-1)}{n}$$

$$D^- = 0,786243$$

Perhitungan nilai selengkapnya seperti pada lampiran B.

4.4.2 Uji Anderson Darling

a. Distribusi Pareto

Uji Anderson-Darling juga akan dilakukan untuk distribusi Pareto dengan menggunakan persamaan (2.27), yaitu :

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \frac{2i-1}{2n+1-2i} \ln \frac{F(x_i)}{1-F(x_i)}$$

maka :

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \frac{2i-1}{2n+1-2i} \ln \left(1 - \frac{k}{x_i}^\alpha \right) + \frac{k}{x_i}^\alpha$$

dengan,

$$i = 1, 2, 3, \dots, n$$

$$n = 1362$$

$$x_i = \text{data ke } i$$

$$\alpha = 0,45233$$

$$k = 3,90$$

sehingga,

$$\begin{aligned} A^2 &= -1362 - \frac{1}{1362} \sum_{i=1}^n \frac{2i-1}{2(1363)+1-2i} \ln \left(1 - \frac{3,90}{x_i}^{0,45233} \right) + \frac{3,90}{x_i}^{0,45233} \\ A^2 &= -1362 - \frac{1}{1362} 2285259 \\ A^2 &= 313,56 \end{aligned}$$

b. Distribusi GEV

Uji Anderson Darling pada distribusi GEV dengan fungsi distribusi kumulatif yang telah diketahui dilakukan sebagai berikut:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \frac{2i-1}{2n+1-2i} \ln \frac{F(x_i)}{1-F(x_i)}$$

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \frac{2i-1}{2n+1-2i} \ln e^{-1+\xi \frac{x_i-\mu}{\sigma}^{\frac{1}{\xi}}} + \frac{1}{\xi}$$

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \frac{2i-1}{2n+1-2i} \ln e^{-1+0,06755 \frac{x_i-30,427}{19,473}^{\frac{1}{0,06755}}} + \frac{1}{0,06755}$$

$$A^2 = -1362 - \frac{1}{1362} 9568,52$$

$$A^2 = 9,4575$$

Nilai uji statistik dari distribusi pareto dan GEV dapat dilihat pada tabel berikut:

Tabel 4.3 Nilai Uji Statistik Kolmogorov Smirnov dan Anderson Darling

No	Distribusi Statistik	Kolmogorov Smirnov	Anderson Darling
1	Pareto	0,37155	313,56
2	GEV	0,05672	9,4575

Nilai uji statistik yang diperoleh menunjukkan bahwa distribusi *Generalized Extreme Value* (GEV) memiliki nilai uji yang minimum dibandingkan dengan distribusi Pareto, sehingga distribusi GEV dapat digunakan untuk memodelkan data pencemaran udara oleh ozon ekstrim di Pekanbaru.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pembahasan sebelumnya diperoleh bahwa distribusi Pareto dan distribusi GEV sesuai dan bisa digunakan untuk memodelkan data pencemaran udara oleh ozon ekstrim di Pekanbaru. Jika dibandingkan dari dua distribusi tersebut maka distribusi GEV lebih sesuai dan lebih baik untuk memodelkan data ozon ekstrim. Hasil uji statistik Kolmogorov Smirnov dan Anderson Darling juga memberikan hasil nilai minimum pada distribusi GEV. Dengan demikian distribusi terbaik dari dua distribusi yaitu distribusi Pareto dan GEV adalah distribusi GEV dengan model sebagai berikut:

$$f(x, \mu, \sigma, \xi) = \frac{1}{19,473} \left(1 + 0,6755 \frac{x-30,427}{19,473} \right)^{\frac{1}{0,6755-1}} e^{- \left(1 + 0,6755 \frac{x-30,427}{19,473} \right)^{\frac{1}{0,6755}}}$$

5.2 Saran

Penelitian ini hanya menggunakan dua distribusi yaitu distribusi Pareto dan distribusi *Generalized Extreme Value (GEV)* dengan hasil distribusi GEV lebih baik dibandingkan dengan distribusi Pareto. Bagi peneliti yang ingin melanjutkan skripsi ini disarankan untuk menggunakan beberapa distribusi yang lain untuk data ekstrim seperti *Generalized Pareto Distribution* dan Distribusi Ekstrim Lainnya.

DAFTAR PUSTAKA

- Choi, B. H., *et. al.* *Distribution Functions of Tsunami Wave Heights. Natural Hazard* 25 : 1-21, 2002.
- Dan'azumi, Salisu, *et. al.* *Modeling the Distribution of Rainfall Intensity Using Hourly Data. American Journal of Environmental Sciences* 6 (3) : 238-243, 2010.
- E Walpole, Ronald dan Raymond H Mayers. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. Bandung : ITB Bandung. 1989.
- Husak, G. J., *et. al.* Use of The Gamma Distribution to Represent Monthly Rainfall in Africa for Drought Monitoring Applications. *International Journal of Climatology*, 27 : 935-944, 2007.
- J Dudewicz, Edward dan Satya N Mishra. *Modern Mathematical Statistics*. John Wiley and Sons, Inc. 1988.
- Koutsoyiannis, Demetris. *On The Appropriateness of The Gumbel Distribution in Modelling Extreme Rainfall. Proceedings of the ESF LESC Exploratory Workshop Held at Bologna, Italy, October 24-25, 2003*.
- Krishnamoorthy, K. *Handbook of Statistical Distributions with Applications*. Chapman & Hall/CRC. 2006.
- Lee, E. T., Wang, J. W. *Statistical Methods for Survival Data Analysis*. 3nd ed. John Wiley & Son, Inc. 2003.
- Mojfeld, H. O. *Forecasting the Heights of Later Wave in Pacific-Wide Tsunamis. Natural Hazard* 22 : 71-89, 2000.
- Pani, Ari. Model Statistik untuk Data Karbon Monoksida (CO). *Prosiding Simposium Kebangsaan Sains Matematik, Fakulti Sains, Universiti Putra Malaysia* : 17, 2009.
- Prager, E. J., *et. al.* *Sains dan Sifat Gempa Bumi, Gunung Berapi, dan Tsunami*. Pakar Karya : Bandung. 2006.
- Thode, H. C. *Testing for Normality*. Marcel Dekker. Inc. 2002.
- Rinne, H. *The Weibull Distribution A Handbook*. Chapman & Hall/CRC. 2009.

Rousas, George. *An Introduction to Probability and Statistical Inference*. Academic Press. 2003.

Sulaiha, Jamaludin, *et. al.* *Fitting Daily Rainfall Amount in Malaysia Using the Normal Transform Distribution*. *Journal of Applied Sciences* 7 (14) : 1880-1886, 2007.

Yendra, Rado, dkk. *Analisis Survival dan Program R*. Yayasan Pusaka Riau : Pekanbaru. 2010.

Zaharim, Azami, *et, al.* *Fitting of Statistical Distributions to Wind Speed Data in Malaysia*. *European Journal of Scientific Research* Vol 1 : pp 6-12, 2009.

Zaharim, Azami, *et, al.* *The Suitability of Statistical Distribution in Fitting Wind Speed Data*. *WSEAS Transactions on Mathematics*, Vol 7, 2008.